A Comparison between Real-Time Polymerase Chain Reaction and Hybrid Capture 2 for Human Papillomavirus DNA Quantitation[1]

Patti E. Gravitt,[2] Robert D. Burk, Attila Lorincz, Rolando Herrero, Allan Hildesheim, Mark E. Sherman, Maria Concepcion Bratti, Ana Cecilia Rodriguez, Kathy J. Helzlsouer, and Mark Schiffman

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892 [P. E. G., A. H., M. E. S., M. S.]; Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205 [P. E. G., K. J. H.]; Albert Einstein College of Medicine, Bronx, New York [R. D. B.]; Digene Diagnostics, Gaithersburg, Maryland [A. L.]; and Proyecto Epidemiologico Guanacaste, Guanacaste, Costa Rica [R. H., M. C. B., A. C. R.]

Abstract

Studies investigating human papillomavirus (HPV) viral load as a risk factor in the development of squamous intraepithelial lesions (SILs) and cancer have often yielded conflicting results. These studies used a variety of HPV viral quantitation assays [including the commercially available hybrid capture 2 (HC 2) assay], which differ in their ability to account for differences in cervical cell collection, linear dynamic range of viral load quantitation, and determination of type-specific *versus* cumulative viral load measures. HPV-16 and HPV-18 viral quantitation using real-time PCR assays were performed to determine whether type-specific viral load measurements that adjust for specimen cellularity result in a different association between viral load and prevalent SIL and cancer, compared with HC 2 quantitation (which does not adjust for cellularity or multiple infections). In general, HPV-16 viral load as measured by real-time PCR increased linearly with increasing grade of SIL while HPV-18 measured using similar techniques increased through low-grade SIL (LSIL), with HPV-18 viral load among high-grade SIL and cancers near the level of cytologically normal women. HC 2 viral load, using the clinical 1.0 pg/ml cut point, differentiated cytologically normal women from women with any level of cytological abnormality (normal *versus* ≥LSIL) but did not change as lesion severity increased. There was no evidence for plateau of HC 2 at high copy numbers, nor was significant variability in total specimen cellularity observed. However, cumulative viral load measurements by HC 2, in the presence of multiple coinfections, overestimated type-specific viral load.

Multiple infections were more common among women with no (32%) or LSIL (51%) [*versus* 23% in high-grade SIL/cancer], partially explaining the lack of a dose response using a cumulative HC2 viral load measure. The nonrandom distribution of multiple infections by case-control status and the apparent differential effect of viral load by genotype warrant caution when using HC 2 measurements to infer viral load associations with SIL and cancer.

Introduction

HPV[3] infection is currently considered to be a necessary but insufficient cause of cervical cancer. Many putative cofactors have already been identified through case-control studies, after controlling for the presence of HPV infection, including both host-specific biological and behavioral factors (*e.g.,* immune competence, cigarette smoking, multiparity, exogenous hormone use, and so on) and virus-specific factors (genotype, variants, viral load; Ref. 1)

With regard to HPV load as a cofactor, multiple case-control studies (2–10) reported that HPV viral load was generally higher among women with cytological abnormalities compared with cytologically normal women. The exact dose response of the relationship was variable, however, and interpretation of the discrepancies is limited by the use of noncomparable and often semiquantitative viral load assays.

The methods used to determine the association of high HPV viral load with prevalent cervical abnormalities include assays based on end point PCR quantitation (2, 3, 5, 7), competitive PCR (11, 12), HC tube (13), and HC 2 (14). All of these methods suffer from a limited quantitative dynamic range of only ~3–4 logs on average. Discrimination of viral load differences outside the linear range of the assay is difficult and leads to misclassification. Furthermore, HC tube/HC 2 and some of the end point PCR assays (2, 3, 5) do not include methods to adjust for differences in the number of cells collected for viral quantitation (specimen cellularity). It is feasible, for example, that without correction for sampling variability, the viral load from two samples with an equivalent proportion of viral genomes/cell but with a significantly different total number of cells/sample will be misclassified as having different viral loads. The resultant viral load misclassification could potentially cause significant bias, especially if cellularity differs between cases and controls as has been shown in at least one study (7).

In addition, HC tube/HC 2 and the competitive PCR of

---

[3] The abbreviations used are: HPV, human papillomavirus; HC, hybrid capture; SIL, squamous intraepithelial lesion; Q-PCR, quantitative PCR; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; LSIL, low-grade SIL; HSIL, high-grade SIL; STM, standard transport medium; OR, odds ratio; IQR, interquartile range; RLU/CO, relative light unit/cutoff value; LRT, likelihood ratio test; ASCUS, atypical squamous cells of undertermined significance.

Caballero *et al.* (12) fail to account for the cumulative effect of multiple infections with more than one HPV genotype in the measurement of viral load. Multiple HPV infections are common (>20% of the total positives) and may be as frequent as 75% of the positive results in HIV-endemic populations (15). If cervical cancer is a clonal disease where >90% of tumors are attributable to a single viral type (16), the contribution of viral quantities from types infecting other sites in the cervicovaginal area are likely to obscure true type-specific viral load contributions to the pathogenesis of each lesion. Furthermore, some studies may find significantly different frequencies of multiple infections between cases and controls (17), which could bias a cross-sectional viral load association with SIL and cancer. HC 2 was designed for use as a qualitative screening assay to identify women with any high-risk HPV infection. Because the assay results are reported as a semiquantitative value, it is inevitable that the cumulative HPV load as measured by HC 2 would be investigated for its ability to additionally stratify women into more refined risk groups based on the quantity of HPV. It would, in fact, be valuable to be able to use a viral load level to differentiate HPV-positive women at high *versus* low risk of prevalent SIL and cancer and thereby increase the specificity of the screening assay. However, one must carefully consider the meaning of a cumulative viral load measure when making etiologic inference for the role of viral load and disease progression under the general one virus–one lesion paradigm.

It is possible that the strength and dose response of HPV viral load associations with prevalent cervical neoplasia and cancer are obscured because of technical differences in the viral quantitation assays. We have examined the likelihood of misclassification of HC 2-derived viral load by comparing HC 2 HPV quantitation with measurement of HPV-16 and HPV-18 viral load using recently described type-specific real-time PCR assays with normalization for total cell number by quantitation of human DNA.[4] HC 2 may be limited by each of the methodological concerns addressed above, whereas the real-time PCR assays used here are type-specific and normalize for differences in specimen cellularity, with a dynamic range of at least 7 logs.[4]

Materials and Methods

Study Population. Cervical cell samples were collected for HPV DNA testing from women enrolling in a population-based, natural history cohort study conducted in Guanacaste, Costa Rica. The design of this cohort study and enrollment data have been described in detail elsewhere (18). Briefly, the Guanacaste Project enrolled 10,049 women from 1993 to 1994 to investigate the role of HPV and cofactors in the development of cervical cancer and to evaluate new cervical cancer screening assays and strategies. Enrolled women answered an interviewer-administered questionnaire (including demographics, reproductive/sexual history, medical history), and 9175 agreed to a pelvic examination [including cells collected by Cervex brush for cytological screening, followed by a Dacron swab in STM (Digene Diagnostics, Gaithersburg, MD) for HPV DNA testing]. Prevalent cervical diagnoses were determined by consensus after expert review of the results from multiple screening tests (including conventional Pap, ThinLayer cytology, PapNet automated

cytology, and cervicography) as previously described (18), with histological confirmation when biopsy was required after colposcopic examination. All enrollment cervical swab samples were tested for HPV using the L1 consensus PCR system and dot blot hybridization for genotyping of 40+ HPV types. HPV-16 and/or HPV-18 enrollment-positive specimens were selected for quantitation via real-time PCR. Written, informed consent was obtained from all participants, and the study protocol was approved by all participating Institutional Review Boards.

A total of 279 women was selected for HPV-16 viral load testing based on a previous L1 consensus PCR-positive result for HPV-16 DNA. Of these 279 samples, 253 (90.7%) were positive by real-time PCR. Of the 26 samples that were negative by real-time PCR, 23 (85%) were of low signal strength in the original L1 PCR assay, indicative of viral loads near the lower limit of sensitivity of PCR. A total of 90 samples was selected for HPV-18 viral load testing based on a positive HPV-18 PCR result using L1 consensus PCR. Eighty-four (93.3%) of these samples were also positive using the quantitative PCR assay. All 6 samples that were HPV-18 negative using the Q-PCR assay were positive with low signal strength by L1 consensus PCR.

The HPV-16- and HPV-18-positive samples selected for viral load analysis do not include an additional 44 HPV-16-positive and 38 HPV-18-positive samples that were identified after retesting 2978 samples by AmpliTaq GOLD (they were originally tested by regular AmpliTaq; Ref. 19). There was not enough material remaining from these samples for viral load determination. Because these samples were missed with the less sensitive AmpliTaq assay, it was assumed that they had viral loads between the lower detection limit of AmpliTaq and the lower detection limit of TaqGOLD (0.08–0.63 HPV-16 copies/20,000 GAPDH and 0.04–0.83 HPV-18 copies/20,000 GAPDH). We therefore imputed the viral load for these samples as the midpoint of this range (0.36 HPV-16 copies/20,000 GAPDH and 0.4 HPV-18 copies/20,000 GAPDH).

The analyses presented here are restricted to the 253 samples with HPV-16-positive and 84 samples with a HPV-18-positive PCR quantitative result. All analyses were repeated including the samples with imputed viral load values, but because the inference was the same, we present the more pure sample group with actual viral load measures for comparison. With the exception of one HPV-16-positive cancer, all specimens gave adequate quantitative GAPDH results allowing for normalization to cell number. This HPV-16-positive cancer was dropped from the present analysis.

HC 2 Testing. HC 2 results were available from a subset of the enrolled women who were included in an analysis to evaluate the efficacy of HC 2 as a cervical cancer screening assay (6). The sampling scheme for that analysis resulted in HC 2 testing on all samples from women with prevalent cytologically abnormal lesions (LSILs, HSILs, and cancer) and on a stratified random sample of women with equivocal and normal cervical diagnoses at enrollment. Of the 253 HPV-16 Q-PCR confirmed positive enrollment samples, 111 (43.9%) had HC 2 results. Of the 84 HPV-18 Q-PCR confirmed positive samples, 28 (33.3%) had HC 2 results, with 3 samples having all three tests (HC 2, HPV-16, and HPV-18). Therefore, 136 samples had both HPV-16 and/or HPV-18 viral load and HC 2 viral load measurements, and these samples are used for direct comparison with HPV PCR quantitation (Fig. 1).

Sample preparation and the HC 2 assay using only the high-risk probe set were performed according to the manufac-

[4] P. E. Gravitt, C. Peyton, C. M. Wheeler, R. J. Apple, R. Higuchi, and K. V. Shah. Reproducibility of HPV 16 and HPV 18 viral load quantitation using TaqMan real-time PCR assays, submitted for publication.
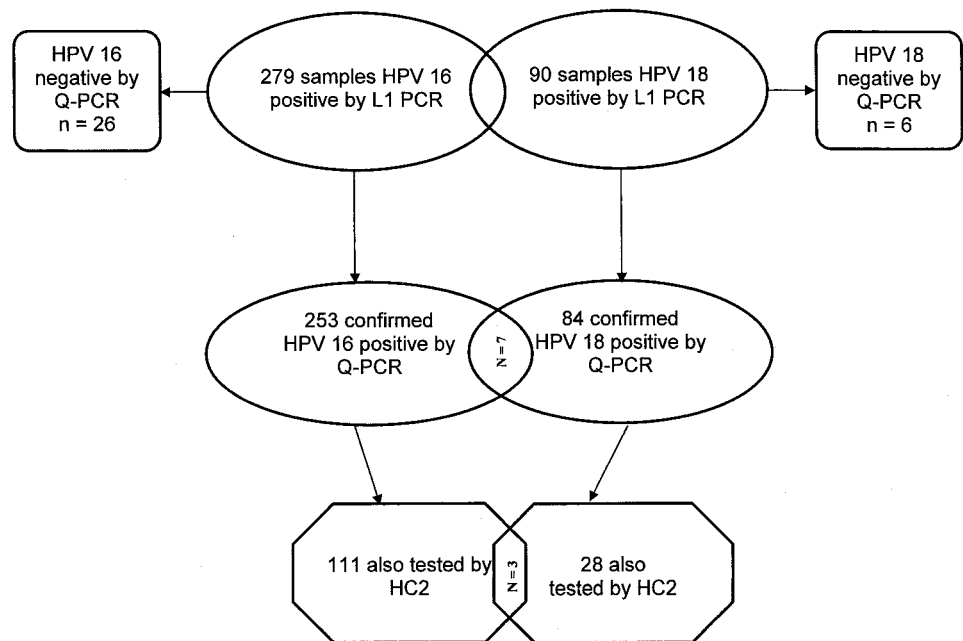
Fig. 1. Sampling schematic for viral load testing based on L1 consensus PCR genotyping result. Overlapping HC 2 sampling indicated in hexagons.

turer's instructions as described previously (6). For each specimen, RLU/CO values were calculated as the ratio of the specimen luminescence relative to the luminescence of the 1.0 pg/ml HPV-16 cutoff standard (~100,000 HPV-16 genomes/ml) and reflect a semiquantitative value of the cumulative viral burden from one or more of 13 oncogenic HPV genotypes (HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68). Any ratio $\geq$ 1.0 RLU/CO using the 1.0 pg/ml standard is considered the clinical cut point for a HPV-positive result (14). Because all samples in this analysis were confirmed as HPV positive by two different PCR systems, we use RLU/CO values < 1.0 in this report for some comparisons with PCR viral load.

HPV PCR Amplification and Genotyping. An aliquot of 400 μl of samples collected into Digene STM (Digene Diagnostics) was added to 100 μl of K buffer containing 400 mg/ml proteinase K and incubated at 55°C for 2 h. Protease was heat inactivated at 95°C for 10 min. A water blank, interspersed after every 20 samples, was processed as a negative control from digestion through PCR. Ten μl of this digest was used for HPV typing, performed as previously described using L1 consensus primer amplification (MY09/11 + HMB01) with genotype discrimination via type-specific oligonucleotide probe hybridization (18), which detected 40+ HPV genotypes, including HPV types 2, 6, 11, 13, 16, 18, 26, 31, 32, 33, 35, 39, 40, 42, 43, 44, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 61, 62, 66, 67, 68, 69, 70, 71, 72, 73, 81, 82, 83, 84, 86, and 3–5 uncharacterized HPVs.

Real-Time PCR Quantitation. HPV-16 and HPV-18 viral load with human cell normalization were measured using the recently described type-specific TaqMan assays (targeting the E6 and E7 open-reading frames, respectively) with normalization for cell number by SYBR-based GAPDH quantitation.[4] These methods were shown to detect $\leq$ 10 HPV genomes/PCR with an analytic specificity > 95%. A 2.5-μl aliquot of DNA that was isolated from Digene STM for HPV L1 PCR amplification, and genotyping was used for HPV PCR viral load determination. The DNA extraction procedure effectively con-

centrates the sample 2-fold; therefore, 2.5 μl of amplification reaction are equivalent to assaying 5 μl (0.5%) of the original sample such that amplification of 0.5% of the 1.0 pg/ml HC 2-positive control (~100,000 genomes/ml) is equivalent to amplification of 500 copies of HPV-16. Therefore, given that for each sample, PCR tests 10-fold less volume relative to HC 2, categorized PCR results of 500-5000 genomes are equivalent to 1–10 RLU/CO by HC 2. Each sample was amplified in duplicate (one undiluted and one 1:10 diluted replicate). Dilution replicates falling within 2.5-fold of the expected 10-fold undiluted:diluted ratio were considered to be sufficient, and the undiluted value was used in the analysis. Dilution replicates falling outside of this range were individually arbitrated, assigning final viral load based on the undiluted sample where low DNA concentrations obscured viral load determination from the diluted sample[4] or based on the diluted sample when there was evidence of PCR inhibition in the undiluted sample. Normalized viral load was then calculated by dividing the HPV copy number by the GAPDH copy number and multiplying by 20,000 for a normalized viral load expressed in units of viral genomes/20,000 cells for convenience.

Statistical Analysis. We sought to estimate the association of prevalent disease status with increasing viral load by each method of viral load measurement. We tested for differences in mean viral load among the different disease categories (*e.g.,* normal, equivocal, LSILs, HSILs, and cancer) by ANOVA. For HPV-16 and HPV-18-specific assays, both mean crude and normalized viral load results were analyzed. Trends in the viral load with increasing disease severity were tested using linear regression, modeling the ordinal categorical disease variable as an independent continuous variable.

To assess whether the differences in viral load and prevalent disease association were attributable to the *a priori* methodological differences between the PCR and HC 2 methods, we performed pairwise comparisons of HPV-16 viral load and HC 2 viral load. To compare the dynamic range of real-time PCR and HC 2, we explored the linear correlation graphically by scatter plots. Spearman's rank correlation was used to assess

Table 1 Summary of viral load measures by enrollment diagnosis

Geometric mean HC 2, HPV-16, and HPV-18 viral load and IQR by each category of cervical abnormality. HC 2 load in regular type represents viral load from all samples with HC 2 testing and in italics the RLU from the HC 2-positive samples only (>1.0 RLU).

Enrollment diagnosis	n	HC2 RLU (pg/ml) Mean (IQR)	n	Normalized HPV-16 viral load copies/20K GAPDH copies (IQR)	n	Normalized HPV-18 viral load copies/20K GAPDH copies (IQR)
Normal	22	14 (2–200)	122	84 (10–582)	56	59 (3–643)
	17	*49 (11–236)*				
Equivocal	6	10 (0.4–74)	29	1206 (134–10,938)	7	985 (49–17,859)
	4	*53 (22–126)*				
LSIL	29	53 (24–275)	32	1315 (164–5570)	7	5597 (114–68,706)
	26	*97 (30–277)*				
HSIL	59	83 (21–332)	55	6325 (2369–21,614)	8	365 (88–821)
	57	*102 (27–332)*				
Cancer	20	83 (10–838)	15	8086 (1550–54,571)	6	659 (258–727)
	19	*106 (12–957)*				
		$P_{trend} = 0.002$		$P_{trend} < 0.001$		$P_{trend} = 0.005$
		$P_{trend} = 0.143$				

relative correlation between the PCR and HC 2 viral load measures. Viral load was categorized in comparable 10-fold increments (see "Materials and Methods," Real-time PCR Quantitation) and categorical agreement measured by κ statistic. We also examined the agreement using a weighted κ statistic to account for minor differences in categorization, where a single category (*i.e.,* 1 log) difference was arbitrarily weighted as 80% agreement, two level differences weighted as 20% agreement, and 3+ level differences weighted as complete disagreement. Asymmetric distribution of discordant results was tested for significance using a test for symmetry (20).

For descriptive comparisons of the effect of normalization for specimen cellularity, we compared the rank of HPV-16 and HPV-18 viral loads from the crude and GAPDH-normalized viral load values using Spearman's rank correlation test. Normalized and crude HPV-16 and HPV-18 viral loads were also categorized by dividing the distribution into thirds and testing the agreement beyond that expected by chance using the κ statistic. Asymmetric distribution of the discordant results was tested for significance using a test for symmetry.

The effect of multiple HC 2-detectable coinfections on the agreement between real-time PCR and HC 2 viral load measurement was explored by comparing the categorized 5 × 5 table (five levels of viral load by HC 2 or PCR) in strata of single (*e.g.,* HPV-16 only) infections and multiple infections. Tests of agreement and discordant symmetry were as described above.

We recategorized the prevalent disease group into a binary outcome, with cases defined as HSIL and cancer diagnoses and controls defined as diagnoses of LSIL, equivocal or normal. These binary categorizations essentially represent women requiring immediate treatment (HSIL and cancer) *versus* clinical management by follow-up (lesions ≤ LSIL). We then modeled the effect of viral load categories from HPV-16 and HPV-18 PCR, as well as HC 2 on this binary independent case outcome variable using logistic regression. The effect of sampling variability (cellularity) on estimate of risk of prevalent disease by viral load was determined by statistical adjustment for ln-GAPDH copy number using multivariate logistic regression. Similarly, the effect of multiple HPV coinfections on the prevalent disease risk estimates by viral load category was determined by statistical adjustment for presence *versus* absence of a second HC 2-detectable genotype using multivariate logistic regression. Full models, including both the viral load and multiple infection or cellularity indicator variables, were compared with the nested viral load model only using likelihood ratio tests

(LRT). Trends in the odds ratio with increasing viral load were tested using logistic regression, where the ordinal viral load categories were modeled as a single, continuous independent variable. Differences were considered statistically significant at $P < 0.05$. All analyses were performed using STATA 7.0 (College Station, TX).

Results

HPV Viral Load by Prevalent Disease Categories. Table 1 summarizes the geometric mean HC 2, HPV-16, and HPV-18 viral load and IQR by each category of cervical abnormality. All three methods of viral load measurement show substantial overlap in range of viral load, particularly when trying to discriminate low- from high-grade neoplasia by viral quantity. The HC 2 data are limited to women with a HC 2 measurement who were selected for the HPV-16/HPV-18 viral load analysis from the PCR data ($n = 136$), including 22 cytologically normal women, 6 with equivocal cytology, 29 LSILs, 59 HSILs, and 20 cancers. There was a significant trend of increasing HC 2 viral load with increasing disease severity when including HC 2 measurements below the 1.0 pg/ml cutoff ($P < 0.002$) but not when restricting the analysis to HC 2 positives based on the 1.0 pg/ml clinical cut point ($P = 0.14$).

Of the 253 samples with a positive HPV-16 viral load test, 122 were cytologically normal, 29 had equivocal cytology, 32 had LSIL, 55 had HSIL, and 15 had a cancer diagnosis at the time the sample for viral load measurement was taken. HPV-16 viral load increased linearly with increasing disease severity after normalization for sample cellularity by dividing the crude HPV viral load by number of GAPDH copies in the specimen ($P_{trend} < 0.001$).

Of the 84 samples with a positive HPV-18 viral load test, 56 were cytologically normal, 7 were graded equivocal, 7 LSILs, 8 HSILs, and 6 as cancers at the time the sample was collected for viral load measurement. The linear increase in HPV-18 viral load occurred only from normal through LSIL diagnoses, dropping to lower viral load in the HSIL and cancer categories.

Because we observed an increasing PCR-derived viral load association with increasing disease severity for HPV-16 but not for HPV-18, the mean HC 2 viral load was recalculated, including the HPV-16-positive samples only ($n = 111$). After restricting to the HPV-16-positive specimens only, the mean HC 2 viral load did not substantially change for the normal,
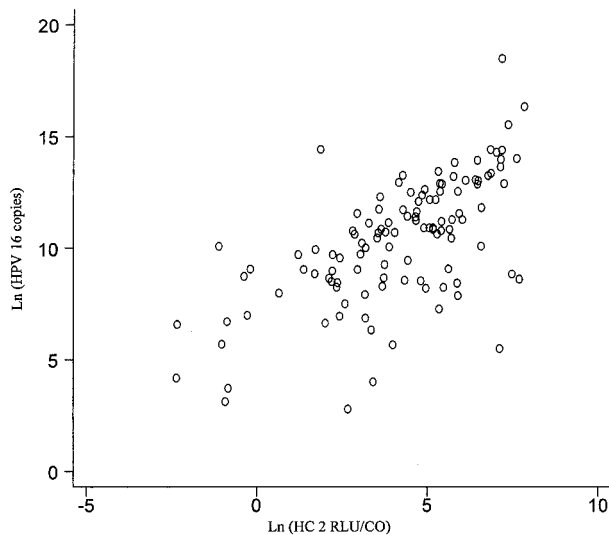
Fig. 2. Correlation of HC 2 and HPV-16 crude viral load.

Table 2 Categorical comparison of HC 2 and HPV-16 real-time PCR

Comparison of crude HPV-16 and HC2 viral load categories among all paired samples (*A*) and stratified by presence of 1 (*B*) *versus* 2+ infections (*C*) with a HC 2-detectable genotype. Concordant numbers are bolded.

	A					
PCR categories[a]	HC 2 categories[a]					Total
	0	1	2	3	4	
0	**4**	0	3	0	1	8
1	3	**3**	8	5	0	19
2	2	6	**16**	6	2	32
3	0	0	9	**25**	1	35
4	0	1	1	7	**8**	17
Total	9	10	37	43	12	111

	B					
PCR categories[a]	HC 2 categories[a]					Total
	0	1	2	3	4	
0	**4**	0	2	0	1	7
1	3	**1**	4	1	0	9
2	1	5	**9**	4	1	20
3	0	0	8	**16**	1	25
4	0	1	1	6	**7**	15
Total	8	7	24	27	10	76

	C					
PCR categories[a]	HC 2 categories[a]					Total
	0	1	2	3	4	
0	**0**	0	1	0	0	1
1	0	**2**	4	4	0	10
2	1	1	**7**	2	1	12
3	0	0	1	**9**	0	10
4	0	0	0	1	**1**	2
Total	1	3	13	16	2	35

[a] 0, <500 HPV-16 PCR copies or <1.0 HC2 RLU; 1, 500–5,000 HPV-16 PCR copies or 1–10 HC2 RLU; 2, 5,000–50,000 HPV-16 PCR copies or 10–100 HC2 RLU; 3, 50,000–500,000 HPV-16 PCR copies or 100–1,000 HC2 RLU; 4, 500,000+ HPV-16 PCR copies or 1,000+ HC2 RLU.

Table 3 GAPDH quantitation by enrollment diagnosis

Back-transformed geometric mean copy number and IQR.

Enrollment diagnosis	n	GAPDH Mean absolute copy number (IQR)
Normal	176	279,246 (166,805–805,840)
Equivocal	35	292,520 (153,803–580,195)
LSIL	38	144,053 (59,219–448,072)
HSIL	61	199,702 (99,032–516,918)
Cancer	20	209,563 (106,184–346,009)
ANOVA		*P* = 0.11

equivocal, and LSIL diagnostic categories (mean HC 2 RLU/CO 19, 19, and 46, respectively). However, the mean HC 2 viral load increased among women with HSIL (RLU/CO = 96) and cancer (RLU/CO = 152), suggesting that the lack of dose response for combined HC 2 load with increased lesion severity was at least partially explained by combining HPV-16- and HPV-18-positive samples in the analysis.

Because we observed some differences between HC 2 and HPV-16 PCR in the cross-sectional association of viral load with disease severity, we sought to determine whether any of the *a priori* methodological differences might explain the discrepancies. We have restricted this comparison to the samples with a HC 2 and HPV-16 PCR test (n = 111); there were too few HPV-18-positive specimens tested by HC 2 to meaningfully compare these assays.

Dynamic Range Differences between HC 2 and HPV-16 PCR. We examined first the direct correlation between HC 2 and HPV-16 real-time PCR graphically using scatterplots of the ln-RLU/CO and ln-HPV-16 copy number (Fig. 2). We saw little evidence of plateau at high copy number, indicating a relatively similar dynamic range between the two assays. The Spearman rank correlation between HC 2 and HPV-16 PCR was 0.62 (P < 0.001), which also indicates relatively good crude agreement between the two viral load measures.

We also examined the agreement by categories of viral load. To meaningfully compare HC 2 and HPV-16 real-time PCR viral load categories, the HC 2 RLU values were converted to genome equivalents (see "Materials and Methods"), and 10-fold categories were compared to assess the linear range of each assay compared with HC 2 results among the total samples tested. Among the total paired samples with HPV-16 viral load measurements (Table 2*A*), the agreement was fair (κ = 0.34; weighted κ = 0.50) with PCR showing higher viral load in 29 samples and HC 2 estimating higher viral load in 26 samples (test for symmetry P = 0.07).

Effect of Normalization on Viral Load Measurement. In this HPV-positive population, there was not a statistically significant difference in mean GAPDH levels (as a correlate of specimen cellularity) between the categories of prevalent disease (Table 3), suggesting that differences in cellularity would not result in a significant differential misclassification bias.

To additionally substantiate the lack of effect of cellularity adjustment on the estimation of viral load in this study, the absolute copy number of HPV-16 viral load estimated via the real-time PCR assays was compared with the normalized HPV-16 viral load. There was excellent rank correlation (Spearman's ρ = 0.89; P < 0.001) between the crude and normalized HPV viral load estimates.

Effect of Multiple HC 2-detectable Infections on Viral Load Estimation. HC 2 viral load measurements encompass a cumulative measure of the amount of HPV DNA/sample from any of the HC 2-detectable genotypes present in that sample (HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, and

Table 4   Association of log increases in viral load by HC 2, HPV-16 PCR, and HPV-18 PCR with prevalent HSIL/cancer

OR = odds ratio; 95% CI = 95% confidence interval.

Viral load	HPV viral load category	Inclusive of all samples with assay-specific testing		Restricted to samples with HC 2 and positive HPV-16 test	
		n	OR (95% confidence interval)	*n*	OR (95% confidence interval)
HC 2 (*n* = 136)	<1 RLU	13	1.0	9	1.0
	1–<10 RLU	16	5.6 (1.1–28.6)	10	3.0 (0.5–19.6)
	10–<100 RLU	43	3.8 (0.9–15.9)	37	1.9 (0.4–8.7)
	100–<1000 RLU	51	7.3 (1.8–30.1)	43	6.6 (1.4–31.3)
	1000+ RLU	13	5.3 (1.0–29.4)	12	4.0 (0.6–25.0)
HPV-16 PCR (*n* = 253)	<500 copies	66	1.0	8	1.0
	500–<5000 copies	49	5.3 (1.1–26.9)	19	1.8 (0.3–11.2)
	5000–<50,000 copies	61	13.4 (3.0–60.7)	32	3.9 (0.7–22.1)
	50,000 copies–<500,000 copies	54	40.0 (8.9–180.4)	35	12.0 (2.0–72.7)
	500,000+ copies	23	41.6 (8.1–212.5)	17	9.8 (1.4–68.8)
HPV-18 PCR (*n* = 84)	<500 copies	30	1.0		
	500–<5000 copies	24	1.3 (0.3–5.8)		
	5000–<50,000 copies	10	2.8 (0.5–15.5)		
	50,000 copies–<500,000 copies	13	1.2 (0.2–7.4)		
	500,000+ copies	7	1.1 (0.1–11.5)		

certain other types because of cross-reactivity); therefore, it is expected that a HC 2-derived viral load will be higher in multiple infections. In the sample set with HPV-16 and HC 2 viral load measurements, 35 of 111 samples (31.5%) were infected with a HC 2-detectable genotype other than HPV-16, as determined by L1 consensus PCR genotyping. We stratified the categorical comparison for HPV-16 and HC 2 in Table 2*A* into a comparison of the HC 2/PCR viral load among women with single HPV-16 infections (Table 2*B*) and women with at least one additional HC 2-detectable genotype coinfection (multiple infection, Table 2*C*).

Among samples with a single HPV-16 infection (*n* = 76), HPV-16 real-time PCR was more likely to estimate a higher viral load relative to HC 2 (*n* = 25) than *vice versa* (*n* = 14). Alternatively, among samples with more than one infection with a HC 2-detectable high-risk type (*n* = 35), HC 2 was more likely to estimate a higher HPV-16 viral load relative to PCR (*n* = 12), with only 4 samples scoring higher by PCR, verifying the cumulative effect of HC 2 viral load measures. Perhaps because of small numbers, these differences did not reach statistical significance. We reanalyzed these data using an expanded definition of multiple infection that included types shown in this population to cross-react with the HC 2 probe pool (HPVs 11, 53, 61, 66, 67, 70, 71, and 81; Ref. 21). Using this modified definition of multiple type infection did not substantially alter the results.

Categorized HPV Viral Load as an Independent Predictor of Prevalent HSIL/Cancer. We also modeled case-control status against a categorized viral load exposure variable to assess whether statistical adjustment for multiple infections or cellularity could improve the agreement between HC 2-based and PCR-based measures of viral load and disease associations. Essentially, we were seeking to determine whether the HC 2 viral load estimates could be combined with the L1 PCR genotyping data to obtain viral load and disease associations (using statistical adjustment) similar to those observed with real-time PCR quantitation. We defined cases as women with prevalent neoplasia that would warrant immediate treatment (HSIL and cancer) and controls as women with LSIL, equivocal, or normal diagnoses at enrollment. The viral load exposure was categorized by log increases in comparable units for the

PCR and HC 2 assays as described in the "Materials and Methods" section to allow for a more direct comparison of the results from the different assays.

The results of these analyses are summarized in Table 4. There is a significant positive association with prevalent HSIL/cancer and a positive HC 2 test result (RLU > 1.0 pg/ml), reflecting the appropriateness of the 1.0 pg/ml cut point suggested for diagnostic purposes. However, no additional increase in OR with increasing RLU values was observed. This is in contrast to an observed increasing strength of association with prevalent HSIL/cancer by log increases in HPV-16 viral load (as measured by PCR) that plateaued at 50,000 copies. No association was seen with HPV-18 viral load and prevalent HSIL/cancer. Because only a subset of the total HPV-16-positive women was tested by HC 2, we restricted the analysis to samples with both HC 2 and HPV-16 viral load results (Table 4, column flush right). The trends were weaker, but the overall pattern was similar, suggesting that the biased selection for HC 2 testing did not completely explain the lack of a dose-response with HC 2 viral load.

To determine the effect of cellularity on the viral load-HSIL/cancer association, we compared the simple viral load models to GAPDH-adjusted models using likelihood ratio tests. Cellularity adjustment did not change the overall trends of association with either HC 2 or HPV-16 log viral load increases, suggesting that cellularity differences were unlikely to explain the HC 2/HPV-16 PCR discrepancies (data not shown). However, adjustment for cellularity significantly improved the HPV-16 PCR viral load model (LRT, *P* = 0.01) but not the HC 2 viral load model (LRT, *P* = 0.24).

We then adjusted for single *versus* multiple (>1 HC 2-detectable genotype) infections to see if statistical adjustment could be used to eliminate the HC 2/PCR discrepancy in assessing HPV type-specific viral load associations. Although adjustment for multiple infection significantly improved the HC 2 viral load model (LRT, *P* < 0.01), the trend was not changed (data not shown). The multiple infection adjustment did not change the trend for the HPV-16 PCR viral load disease association (data not shown), nor was the model significantly improved by this adjustment (LRT, *P* = 0.16).

Discussion

We confirmed that HPV viral load was higher among HPV-16- and/or HPV-18-positive women with any prevalent cervical abnormality compared with cytologically normal women, independent of the assay used for viral load determination. However, the association between viral load and lesion severity varied by assay and genotype.

HPV-16 viral load, as measured by real-time PCR, was associated with increasing severity of cervical neoplasia, even after adjustment for specimen cellularity. The main effect of adjustment for cellularity was a relative decrease in the HPV-16 viral load among the 30 women with equivocal cytological diagnoses who had somewhat higher cell equivalents/unit volume. In contrast to earlier reports, the mean cell concentration/sample did not vary significantly between lesion grades, although the generalizability of our observations is limited to HPV-16 and/or HPV-18-infected women (7). The explanation for this discrepancy between our results and those reported previously is unclear. Although Swan *et al.* (7) used end point quantitation and corrected for cellularity by measurement of the β-globin gene in their study, the assay differences are unlikely to have caused the observed linear effect if one did not exist. The samples for this analysis were collected using standardized protocols by the same clinicians (directed to the cervical overall survival), which is consistent with the low variability of cell collection observed here, something that is more difficult to control in other settings (*e.g.,* multisite studies or self-collection). In addition, the sample tested for viral load was collected after a Pap smear sample by cervical brush, which could have minimized the variability of sample collection on the second swab. Alternatively, because the majority of women in Costa Rica were found to have high levels of cervical inflammation at the time the swab was collected (22), the contribution of human DNA from inflammatory rather than epithelial cells may have biased this association to the null. Although we did not observe differences in cellularity by lesion grade, each study would be wise to assess the extent of variability of cell collection and adjust when necessary. Additionally, methodological developments that would allow specific quantitation of epithelial cells and proportions of infected *versus* uninfected cells would be of great value in elucidating the effect of viral load *in vivo* on cervical carcinogenesis.

HPV-18 viral load increased through LSIL but was only modestly higher among HSIL and cancers compared with cytologically normal women. This effect persisted after adjusting for variability in sample collection. The contrast in association of HPV-16 *versus* HPV-18 viral load and prevalent disease is consistent with that reported earlier (7). The HPV-18 viral load values in our study were log normally distributed with a similar range to HPV-16, and the assay was previously shown to have good amplification efficiency and reproducibility,[4] suggesting that the lack of association between HPV-18 viral load and prevalent neoplasia is not because of assay-specific misclassification. It will be interesting to determine whether the decrease in HPV-18 viral load in HSIL and cancer may be because of increased frequency of integration or localization of HPV 18-associated lesions to the endocervix, which is less efficiently sampled compared with the ectocervical lesions that may be more likely to harbor HPV-16 infection. Future studies will test these possibilities.

HC 2 results in this study are from a convenience sample of women selected into the HPV-16/HPV-18 viral load cohort study who happened to have had a HC 2 measurement from the same swab. HC 2 was performed in the Guanacaste Project only for a nested screening efficacy study, which tested nearly all prevalent lesions LSIL or greater, plus a stratified random sample of cytologically normal women (sampling schema detailed in Ref. 6). The sampling for HC 2 testing among normal women attempted to capture all HPV-positive participants, but HPV prevalence was only determined at that point by the relatively insensitive first generation HC tube test. Therefore, the HC 2 convenience sample is likely biased to HPV infections with higher viral load, particularly among the cytologically normal women. This selection bias may explain the attenuated trend of HC 2 viral load between cervical disease *versus* normals. However, restriction of the HPV-16 analysis to women with HC 2 measures resulted in similar patterns relative to the analysis that includes all women with HPV-16 infection, suggesting that the effect of selection bias is not large. Additionally, HC 2 viral load disease associations should be interpreted here as the HC 2 viral load association with prevalent neoplasia among HPV-16- or HPV-18-positive women. Given these caveats, we still confirmed that the combined HC 2 viral load was higher among women with any cytological abnormalities, as was seen in other studies (10). Among women with a HC 2 result > 1.0 RLU (the clinical cutoff), viral load did not discriminate the degree of lesion severity (*i.e.,* LSIL *versus* HSIL), results that are consistent with a recent large randomized triage study of ASCUS and LSIL referrals (10). This may be because of the combined effect of several factors. First, because LSIL samples have similar viral load to the HSIL/cancers by HC 2, inclusion of LSIL in a control definition is likely to reduce the OR when modeling a viral load association with true precancerous and cancerous lesions (*i.e.,* HSIL and cancer). Furthermore, the higher viral load among the equivocal/LSIL diagnoses may be partially explained by an increase in the proportion of infections with more than one HC 2-detectable genotype among women with LSIL (51%) relative to either cytologically normal (32%) or HSIL and cancers (23%). Because HC 2 is a cumulative measure of all high-risk genotypes simultaneously sampled from the cervix, the differential proportion of multiple infections by diagnostic category reflects an important misclassification bias when attempting to estimate the etiologic role of a single HPV type-specific viral load in lesion development. Also, we observed a significant HPV-16 viral load association by PCR but no HPV-18 viral load association using similar methods. If viral load associations differ by genotype, as suggested by these and other (7) observations, the cumulative HC 2 viral load may be misleading.

When restricted to HPV-16 single infections, the correlation of viral load measures by HC 2 and PCR was good. We saw linearity and good correlation of HC 2 across the entire dynamic range of values measured with no evidence for plateau at higher copy numbers. However, we did find 8% of the HPV-16-positive samples and 14% of the HPV-18-positive samples to fall below the 1.0 pg/ml cut point used in standard HC 2 clinical applications. These samples encompass the lower 2 logs of PCR quantitation. To the extent that differentiation within this range may be important in predicting disease prevalence or progression, use of HC 2 could be limited. However, there is some question regarding the validity of positive tests with viral load < 100 copies/test in that chance contamination could be mistaken for true infection, leading to inflated OR of viral load and neoplasia. Given that the samples falling into the lower 2-log range in this study were confirmed by a second PCR assay, it is unlikely that the lower signal represented contamination, unless the samples were contaminated at collection or during specimen processing (also unlikely given the rigorous quality control used in this study).

In summary, the difference between type-specific and cumulative viral load measurements appears to have an effect on the strength and dose-response of the viral load association with prevalent neoplasia. When estimating risk of prevalent or incident cervical neoplasia as a function of HPV viral load, the effect of using a combined viral load measure can lead to misclassification and attenuation of the risk estimates. Because the extent of misclassification will depend on the distribution of multiple infections and genotypes among the diagnostic groups in each population, this effect might explain some of the discrepant observations in the literature where HC 2 was used. Because the analytic performance of HC 2 and HPV-16 PCR was comparable among single infections, there may be validity in using HC 2 for viral load estimation in single infections. However, caution is recommended in generalizing results from this approach, as substantial bias could be introduced when excluding women with multiple infections from an analysis.

When investigating an etiologic role of viral load in the progression of cervical neoplasia, the viral load measurement must do more than serve as a marker, it must adequately reflect the quantity of virus in the cervix at that sampling. Because of limitations in collecting a truly random and representative sample of the entire cervix (*e.g.,* testing predominately surface *versus* basal epithelial cells), we may never fully understand the role of viral load in cervical cancer. However, use of quantitative PCR methods may better represent the viral load *in vivo* and offers an improved assay for determining the role of viral load in the etiology of cervical cancer. Verification of type-specific differences in viral load and HSIL/cancer associations observed in this study is required and will be an important next step in understanding the potential heterogeneity of papillomavirus-induced cancers.

References

1. Bosch, F. X., Lorincz, A., Munoz, N., Meijer, C. J., and Shah, K. V. The causal relation between human papillomavirus and cervical cancer. J. Clin. Pathol. (Lond.), *55:* 244–265, 2002.

2. Bavin, P., Giles, J., Deery, A., Crow, J., Griffiths, P., Emery, V., and Walker, P. Use of semi-quantitative PCR for human papillomavirus DNA type 16 to identify women with high grade cervical disease in a population presenting with a mildly dyskaryotic smear report. Br. J. Cancer, *67:* 602–605, 1993.

3. Flannelly, G., Jiang, G., Anderson, W., Melvin, W., Mann, E., and Kitchener, H. Serial quantitation of HPV-16 in the smears of women with mild and moderate dyskaryosis. J. Med. Virol., *47:* 6–9, 1995.

4. Cope, J., Hildesheim, A., Schiffman, M., Manos, M., Lorincz, A., Burk, R., Glass, A., Greer, C., Buckland, J., Helgesen, K., Scott, D., Sherman, M., Kurman, R., and Liaw, K-L. Comparison of the hybrid capture tube test and PCR for detection of human papillomavirus DNA in cervical specimens. J. Clin. Microbiol., *35:* 2262–2265, 1997.

5. Forslund, O., Lindqvist, P., Haadem, K., Czegledy, J., and Hansson, B. HPV 16 DNA and mRNA in cervical brush samples quantified by PCR and microwell hybridization. J. Virol. Methods, *69:* 209–222, 1997.

6. Schiffman, M., Herrero, R., Hildesheim, A., Sherman, M. E., Bratti, M., Wacholder, S., Alfaro, M., Hutchinson, M., Morales, J., Greenberg, M. D., and Lorincz, A. T. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. J. Am. Med. Assoc., *283:* 87–93, 2000.

7. Swan, D., Tucker, R., Tortolero-Luna, G., Mitchell, M., Wideroff, L., Unger, E., Nisenbaum, R., Reeves, W., and Icenogle, J. Human papillomavirus (HPV) DNA copy number is dependent on grade of cervical disease and HPV type. J. Clin. Microbiol., *37:* 1030–1034, 1999.

8. Sun, C. A., Lai, H. C., Chang, C. C., Neih, S., Yu, C. P., and Chu, T. Y. The significance of human papillomavirus viral load in prediction of histologic severity and size of squamous intraepithelial lesions of uterine cervix. Gynecol. Oncol., *83:* 95–99, 2001.

9. Healey, S. M., Aronson, K. J., Mao, Y., Schlecht, N. F., Mery, L. S., Ferenczy, A., and Franco, E. L. Oncogenic human papillomavirus infection and cervical lesions in aboriginal women of Nunavut. Canada. Sex Transm. Dis., *28:* 694–700, 2001.

10. Sherman, M. E., Schiffman, M., and Cox, J. T. Effects of age and human papilloma viral load on colposcopy triage: data from the Randomized Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study (ALTS). J. Natl. Cancer Inst. (Bethesda), *94:* 102–107, 2002.

11. Jacobs, M., Walboomers, J., van Beek, J., Voorhorst, F., Verheijen, R., Meijer, C., van den Brule, A., Helmerhorst, T., and Snijders, P. A quantitative polymerase chain reaction-enzyme immunoassay for accurate measurements of human papillomavirus type 16 DNA levels in cervical scrapings. Br. J. Cancer, *81:* 114–121, 1999.

12. Caballero, O., Villa, L., and Simpson, A. Low stringency-PCR (LS-PCR) allows entirely internally standardized DNA quantitation. Nucleic Acids Res., *23:* 192–193, 1995.

13. Cox, J. T., Lorincz, A. T., Schiffman, M. H., Sherman, M. E., Cullen, A., and Kurman, R. J. Human papillomavirus testing by hybrid capture appears to be useful in triaging women with a cytologic diagnosis of atypical squamous cells of undetermined significance. Am. J. Obstet. Gynecol., *172:* 946–954, 1995.

14. Lorincz, A. T., and Anthony, J. Advances in HPV detection by hybrid capture. Papillomavirus Report, *12:* 145–154, 2001.

15. Gravitt, P. E., Kamath, A. M., Gaffikin, L., Chirenje, Z. M., Womack, S., and Shah, K. V. Human papillomavirus genotype prevalence in high-grade squamous intraepithelial lesions and colposcopically normal women from Zimbabwe. Int. J. Cancer, *100:* 729–732, 2002.

16. Stoler, M. H., and Baber, G. The frequency of multiple HPV infection in cervical neoplasia. 19th International Papillomavirus Conference, 9-1-2001. Florianopolis, Brazil, 2001.

17. Fife, K. H., Cramer, H. M., Schroeder, J. M., and Brown, D. R. Detection of multiple human papillomavirus types in the lower genital tract correlates with cervical dysplasia. J. Med. Virol., *64:* 550–559, 2001.

18. Herrero, R., Schiffman, M. H., Bratti, C., Hildesheim, A., Balmaceda, I., Sherman, M. E., Greenberg, M., Cardenas, F., Gomez, V., Helgesen, K., Morales, J., Hutchinson, M., Mango, L., Alfaro, M., Potischman, N. W., Wacholder, S., Swanson, C., and Brinton, L. A. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. Rev. Panam. Salud Publica, *1:* 362–375, 1997.

19. Castle, P. E., Schiffman, M., Gravitt, P. E., Kendall, H., Fishman, S., Dong, H., Hildesheim, A., Herrero, R., Bratti, M. C., Sherman, M. E., Lorincz, A., Schussler, J. E., and Burk, R. D. Comparisons of HPV DNA detection by MY09/11 PCR methods. J. Med. Virol., *68:* 417–423, 2002.

20. Bowker, A. H. A test for symmetry in contingency tables. J. Am. Stat. Assoc., *43:* 572–574, 1948.

21. Castle, P. E., Schiffman, M., Burk, R. D., Wacholder, S., Hildesheim, A., Herrero, R., Bratti, M. C., Sherman, M. E., and Lorincz, A. Restricted cross-reactivity of Hybrid Capture 2 with non-oncogenic human papillomavirus types. Cancer Epidemiol. Biomark. Prev., *11:* 1394–1399, 2002.

22. Castle, P. E., Hillier, S. L., Rabe, L. K., Hildesheim, A., Herrero, R., Bratti, M. C., Sherman, M. E., Burk, R. D., Rodriguez, A. C., Alfaro, M., Hutchinson, M. L., Morales, J., and Schiffman, M. An association of cervical inflammation with high-grade cervical neoplasia in women infected with oncogenic human papillomavirus (HPV). Cancer Epidemiol. Biomark. Prev., *10:* 1021–1027, 2001.